COMMUNICATION unBREAKDOWN

Finding Targets for Earnest Conversation and Its Application to Voting

Katie

Nick

Dani

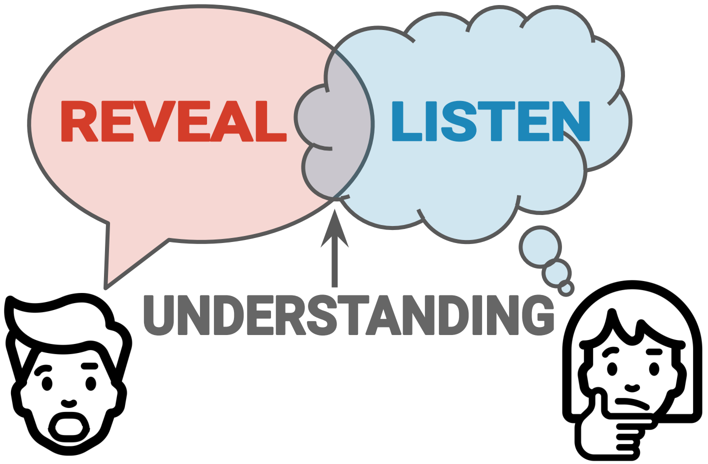
Kevin

Image created in Adobe Illustrator under the tutelage of [Andrei Stefan]
Communication Breakdown song by [Led Zeppelin]

People’s Action is changing hearts and minds one conversation at a time.

PEOPLE’S ACTION

DEEP CANVASSING



Icons by Symbolon courtesy of the Noun Project

Deep Canvassing is unique in how much listening is involved.

Traditional Canvass

- Canvasser hurriedly reads from a script
- Canvasser delivers a "message"
- Usually under 30 seconds
- Canvassers tell voters what to think

Deep Canvass

- Voter does more talking than canvasser
- Voter candidly describes personal experiences
- Upwards of 30-60 minutes
- Voters draw own conclusions

Icons by Gan Khoon Lay courtesy of the Noun Project

It is becoming increasingly difficult and more expensive to reach target voters.

The Plan

8M phone calls

24K volunteers

8 states

The Challenge

3.1% change opinion

$14.72 per conversation

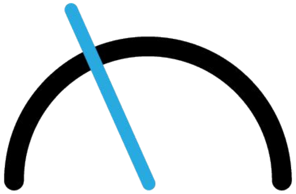
$475 per change in opinion

The Opportunity

- ⇧ Improve conflicted voter targeting
- ⇧ Increase number of positive conversations

People’s Action connected with 33% more voters using our machine-learning informed method.

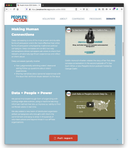
33%



Increase in Target Voters

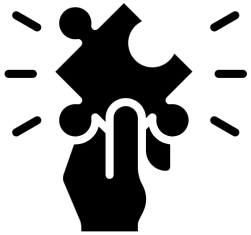
Data + People = Power

“We paired this breakthrough form of organizing with cutting-edge data science, **using a machine-learning-informed method** that lets us improve our ability to find conflicted voters in real time.”*



*From People’s Action website
https://peoplesaction.org/peoples-action-2020/

We delivered our solution by defining a process, creating an MVP, and deploying our models.



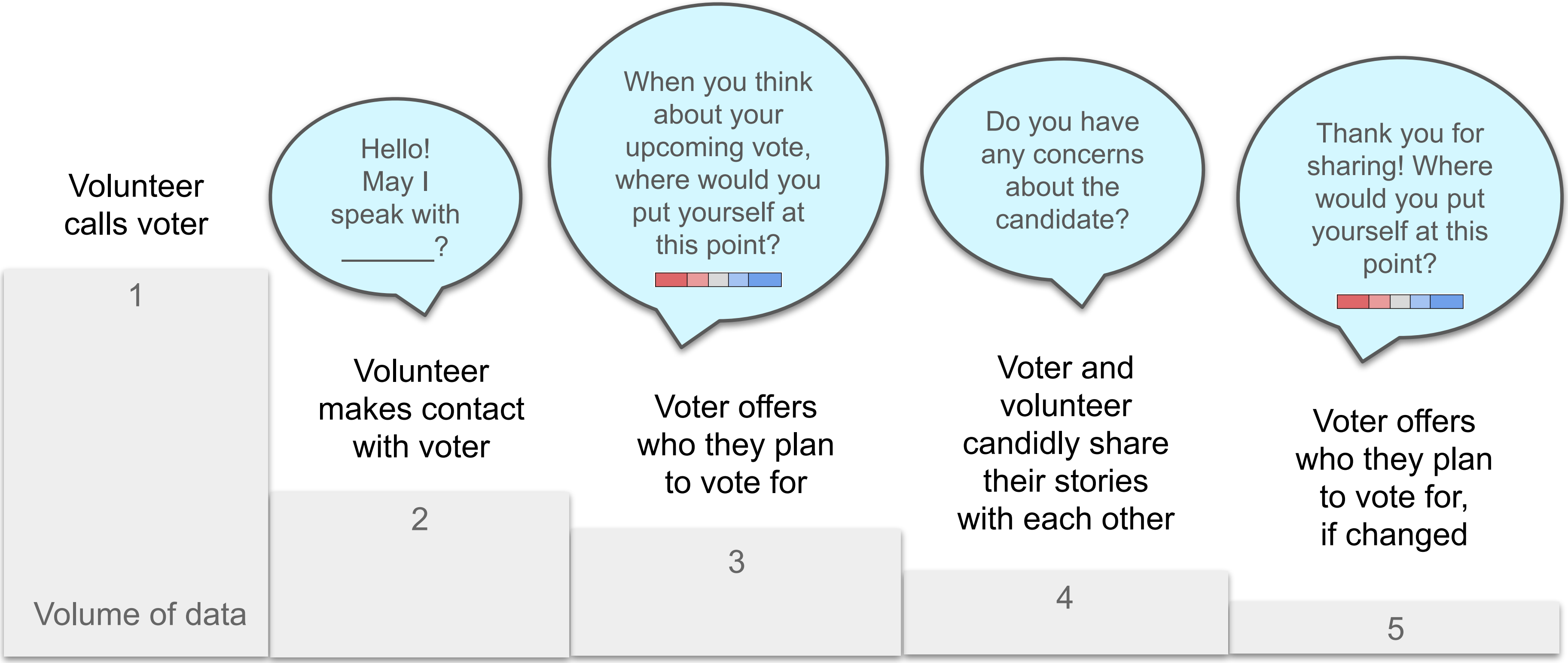
The Process

The MVP

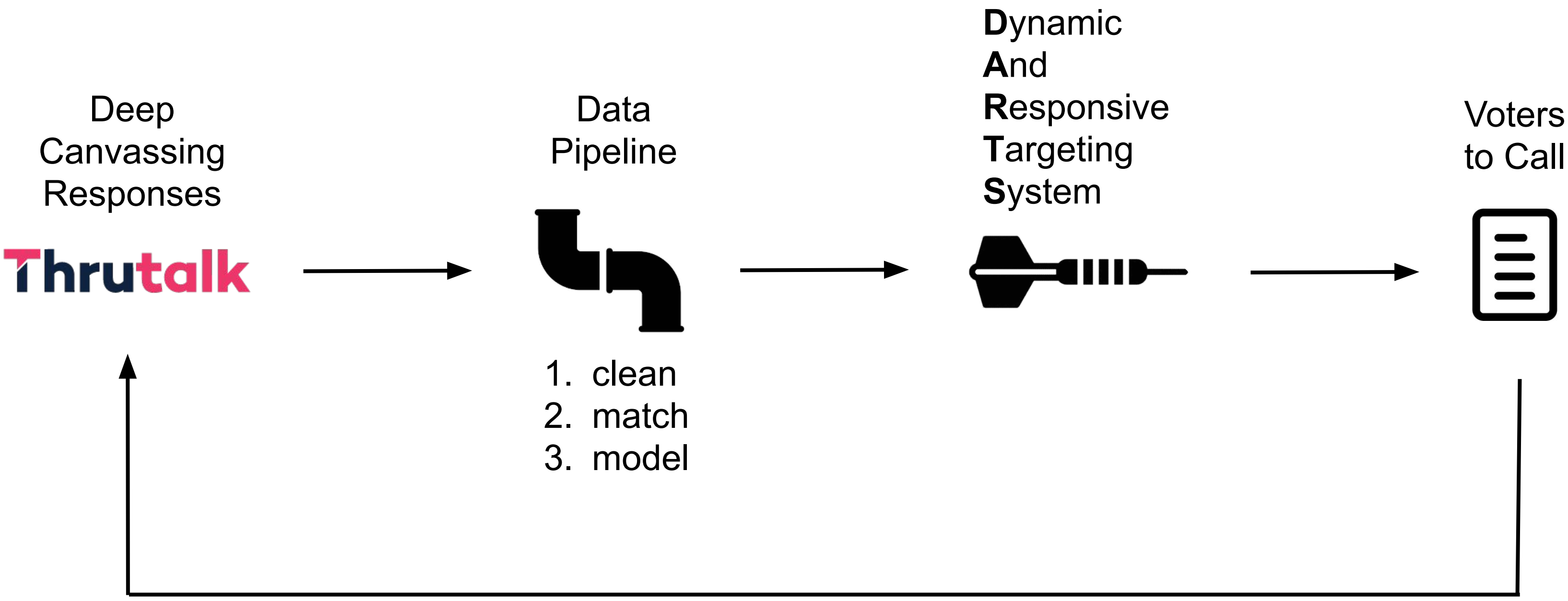
$$f(x_i)$$

The Models

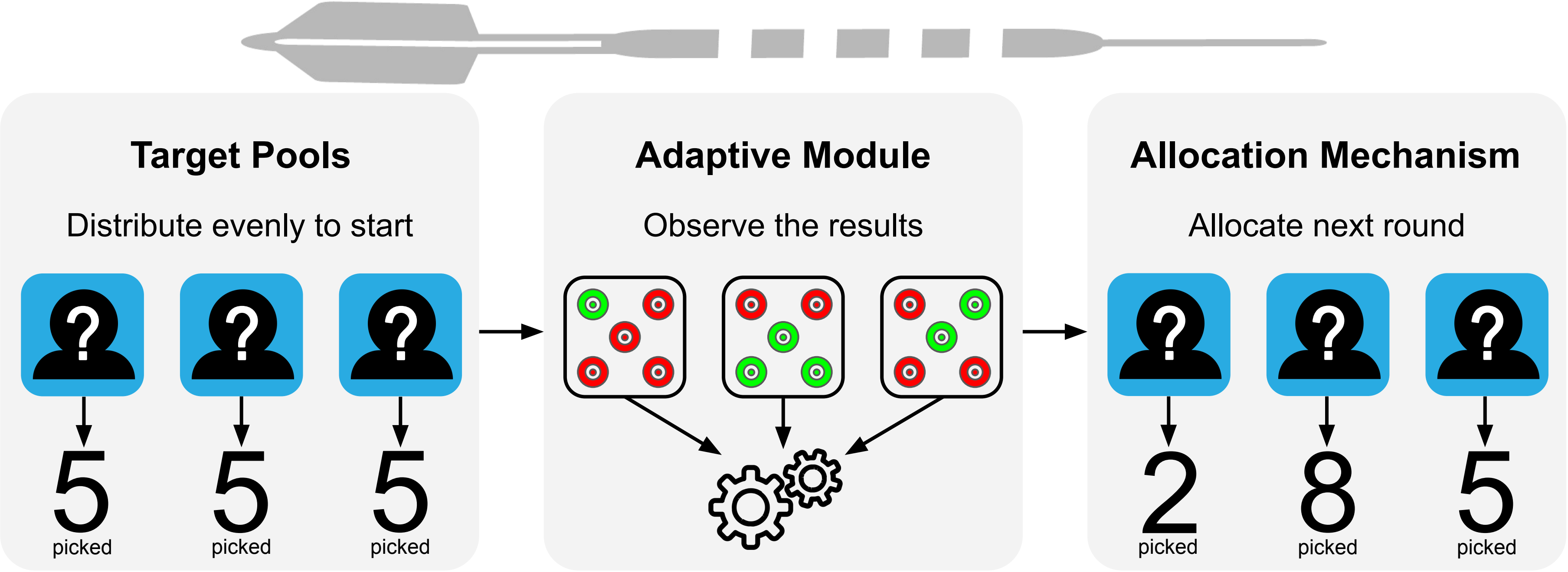
Deep Canvassing is about sharing stories.



We identify targets through an active feedback loop.



Finding quality targets requires a dynamic system.



- Targets share attributes
- Unknown, dynamic distribution
- Flexible implementation

- Adapts from history
- Dictates next round
- Multi-armed Bandit

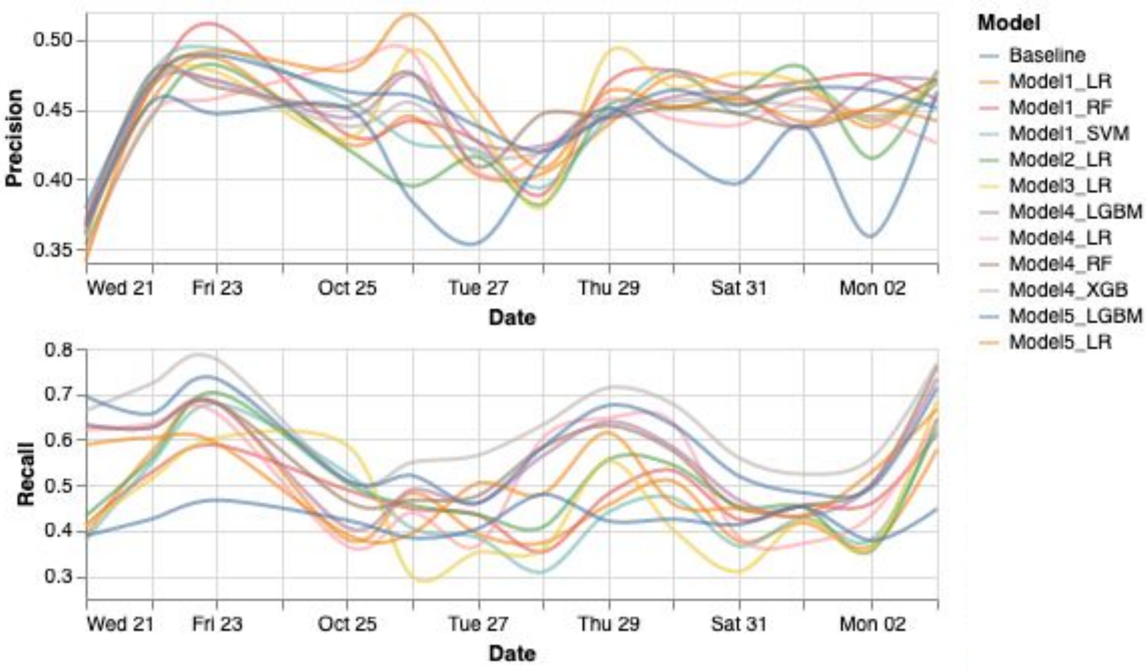
- Adaptive Module as input
- Picking strategy
- Domain dependent

Icons by Tesyar Azhari courtesy of the Noun Project

In dynamic domains, concept drift becomes concept shift.

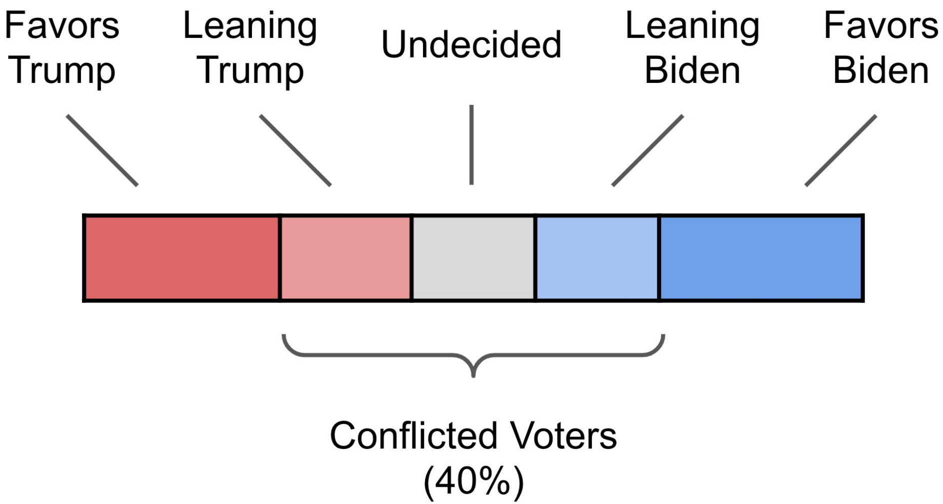
- Rapid concept drift:
  - Shifts in political climate
  - Noisy data
  - Human psychology

- MVP developed to account for:
  - Benefits of model diversification
  - Unforeseen uncertainty
  - Better predictive strategy



Precision and Recall scores for models over test period

We trained our models on live campaign data.

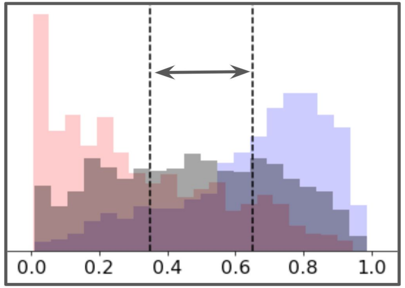


Features

- Demographics
- Socio-economic status
- Jurisdiction
- Voting history
- Moral foundation theory
- Survey scores on political and social issues

We diversified our options by pursuing four strategies.

Strategy 1



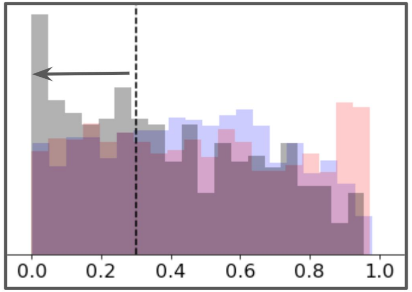
Target mid-range probabilities

Training labels

Class 1	0	0	1	1

Uncertain binary classification model

Strategy 2



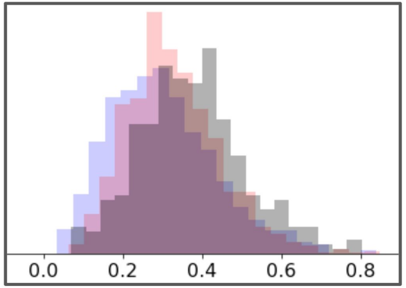
Target low difference in probabilities

Training labels

Class 1	0	0	0	1	1
Class 2	1	1	0	0	0

Difference between two binary classification models

Strategy 3



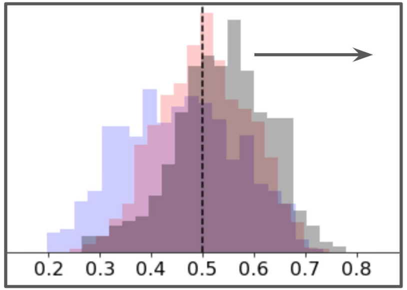
Target undecided class

Training labels

Class 1	0	0	1	2	2

Multi-class classification model

Strategy 4



Target high probabilities

Training labels

Class 1	0	1	1	1	0

Single binary classification model

We trained our models on a variety of algorithms.

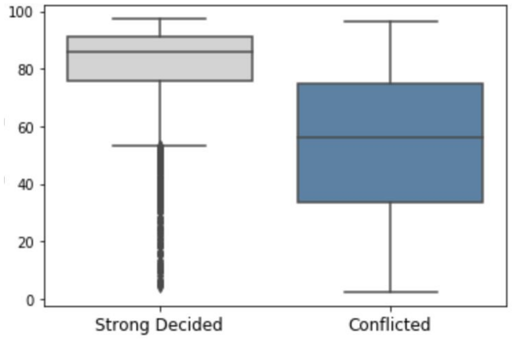
Algorithm	Strategy	Pros	Cons
Logistic Regression	1, 2, 3, 4	**Generalizes well for a variety of classification problems, even with imbalanced classes; high explainability**	May underperform if features are highly correlated/not linear
Random Forest	1, 4	**Builds trees iteratively using features as split points; avoids overfitting through bagging; high explainability**	Underperforms with imbalanced classes
SVM	1	Effective in high-dimensional space; good explainability	Slow to implement; may underperform depending on kernel function
KNN	1, 4	Tends to fit training data very well; difficult to explain	**Overfits and does not generalize well**
AdaBoost	4	Increases weights on observations that are difficult to predict in each iteration; performs well with imbalanced data; high explainability	**Tends to underperform with high degree of noise in the data**
XGBoost	1, 4	Efficient boosted decision trees classifier; generalizes well; high explainability	Underperforms with imbalanced classes
LightGBM	4	**Very fast execution of boosted decision trees classifier; generalizes well; high explainability; better with Boruta**	Underperforms with imbalanced classes

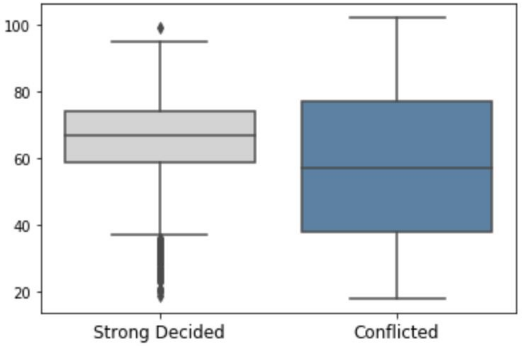
Recall was an important metric for evaluation.

Strategy / Algorithm	Recall	Precision	Top 3 Features
Baseline: Logistic Regression	42%	41%	N/A
Strategy 1: Random Forest	46%	46%	Immigration Sentiment, Gun License, DACA Sentiment
Strategy 2: Logistic Regression	48%	45%	Immigration Sentiment, Gun License, Age
Strategy 3: Logistic Regression	42%	47%	Age, Sanctuary City Sentiment, Children Present
Strategy 4: Light GBM	**60%**	**50%**	Election Turnout, Tea Party, Affirmative Action Sentiment

Model explainability was important for our stakeholder.

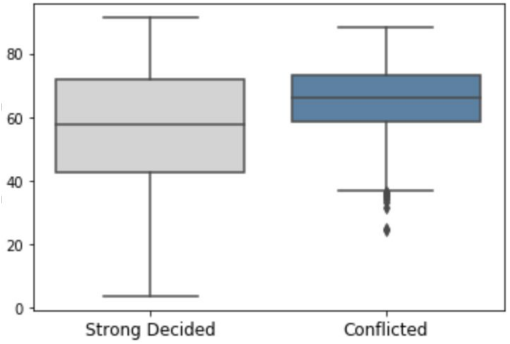
Election Turnout score
for Strategy 4
XGBoost

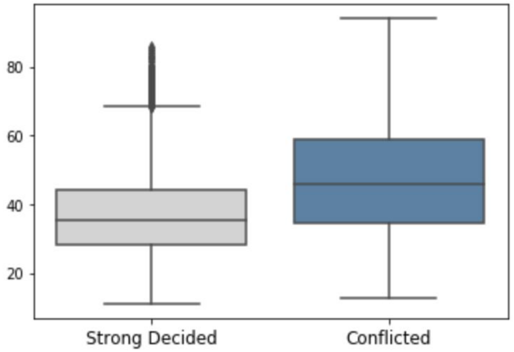




Age
for Strategy 4
XGBoost

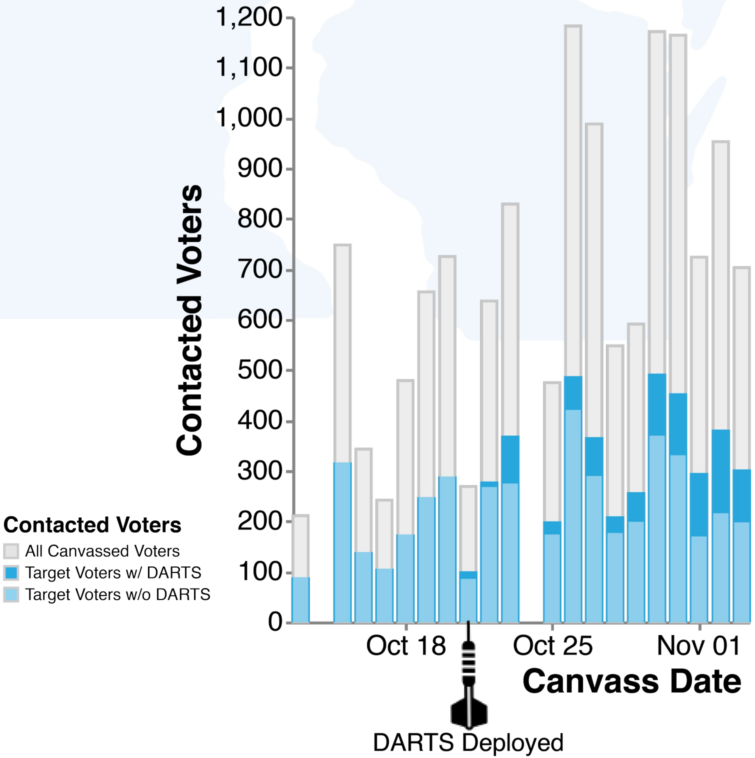
Immigration score
for Strategy 3
Logistic Regression





Universalism score
for Strategy 4
LightGBM

DARTS improved targeting for People's Action by 33%.

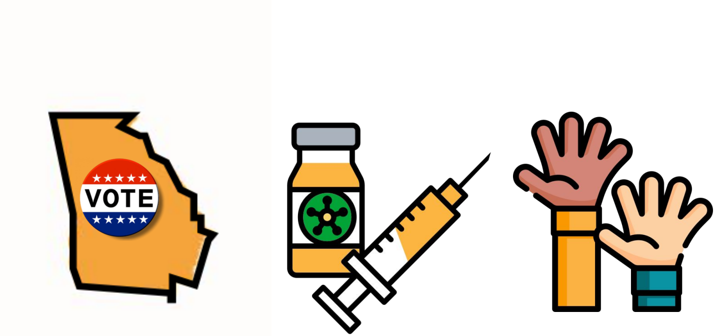


Key Learnings

- Mission is impactful
- Deep Canvassing is challenging
- Matching voter demographic data is not straightforward
- Active feedback loop from the campaign requires intensive data collaboration
- DARTS is an effective way to solve similar problems

Our work can be applied to other big problems.

- Finding conflicted voters for state elections
- Finding conflicted people about vaccines
- Finding volunteers, change agents and donors



To find out more:

Install our software	Fork us on Github	Visit our website
python™ Package Index		
https://pypi.org/project/darts-berkeley/	https://github.com/Berkeley-Data/darts	https://unbreak.info


COMMUNICATION un BREAKDOWN
THANK YOU